

### REMARKS

In the outstanding Office Action, claims 40-59 were presented for examination. Claims 40-47 and 57-59 were rejected under 35 U.S.C. §102(e) as being anticipated by Minematsu (US 6,249,763) and claims 48-56 were rejected under U.S.C. §103 as being unpatentable over Minematsu and further in view of applicant's admitted prior art.

#### *Request for Entry under 37 CFR §1.116*

The Office Action has been most carefully studied. In this amendment no claims are added and no new issues are raised. Furthermore, the amendments are believed to place the application in condition for allowance or, in the alternative, in better condition for consideration on appeal. Accordingly, entry of this amendment under 37 CFR §1.116 is respectfully requested.

#### *Interview*

Examiner Azad's courtesy in arranging and conducting a telephone interview with the undersigned, on September 13, 2005, is greatly appreciated by applicant. During that interview, proposals for amendment of base claim 40, along the lines of the amendments now presented, were discussed. Various merits of applicant's claimed invention and some shortcomings of the reference to Minematsu were also discussed. Examiner Azad recognized that phrase matching and recognition were not disclosed in the reference to Minematsu and kindly indicated that he would consider a further written submission with an open mind.

#### *Amended Claims*

Base claims 40 and 57 have been amended to make explicit subject matter that was inherent in the respective claim before amendment, and for readability. The amendment to clause ii) of claim 40 and correspondingly in claim 57, finds support in the specification at page 27, lines 19-20 which language was implicit in the usage of the

term "phrase" in the claims prior to amendment. Claim 42 has been amended, without narrowing, to be drawn to what was an optional limb of the claim prior to amendment, namely, digitizing, analyzing and storing the information generated by the speaker. Minor corrections have been made to the syntax of claim 42, for readability.

### *The Claimed Invention*

New claim 40 is directed to the invention as it relates to a computerized speech recognition method which employs a database of digital representations, and associated graphic representations, of properly pronounced known spoken phrases. This database is employed to make comparisons with audible sound components of input speech and to identify matches with properly pronounced spoken phrases. Input spoken phrases are matched with properly spoken phrase models in the phrase model database, yielding matched phrases.

As is required by claim 42; the properly pronounced spoken phrases in the database can be generated by speech-trained speakers.

As is further explained below, the proper pronunciation of a phrase may differ from that of the words, or word parts, that compose the phrase.

The significance of phrase pronunciation can be better understood from the following example. The normal continuous speech pronunciation of the phrase "keep back" is not that of the sequential pronunciation of the two words "keep" and "back" but is more typically correctly pronounced almost as though the adjacent cognate consonants "p" and "b" were a single consonant having a sound similar to each letter, albeit a little different from each. Correct pronunciation in fluent speech may be closer to "keeback" than to "keep" followed by "back". Yet, "keeback" will not appear in a word database derived from one or more general purpose dictionaries and therefore is not

system-recognized. Proper fluent pronunciation of many phrases eliminates the audible separation that occurs between properly spoken words, running the words together to create a new sound which is distinct from the sounds yielded by sequential pronunciation of the words. For example, "bad actor" is usually pronounced "badactor" rather than "bad" "actor".

Absent the phrase matching of the claimed invention, such properly pronounced phrases may be incorrectly identified as another word or words.

Applicant's claimed invention enables this problem to be addressed by providing a database of properly pronounced phrases as well as words, which phrases desirably are generated by trained speakers. Minematsu is unaware of the problem and has no provision for discriminating between the different pronunciations of a properly pronounced phrase and of the words in the phrase. For example, Minematsu's system might identify the "kee" component of "keeback" as an error.

By avoiding such errors, the invention as claimed improves recognition accuracy. Because phrases represent larger speech components than do their component words or word parts, processing speed is enhanced. There is no suggestion in Minematsu or any other art known to applicant, that these benefits might be obtained by matching input speech components with digital representations of properly pronounced phrases.

#### *Minematsu*

Minematsu's objective is to recognize improper pronunciation from a foreign speaker as though it were correct pronunciation. To do this, Minematsu provides "analogous words" alternatives to "candidate words" for the system to consider. The system must then employ syntactic parsing or user input to select between the alternatives. Unlike applicant's claimed invention, nothing about Minematsu's

disclosure will assist, or is intended to assist, a system to more accurately or rapidly recognize properly spoken language. Minematsu's objective is better recognition of English as it is often spoken by a Japanese speaker. While this objective is presumed to be achievable by Minematsu's disclosed system, the need to select between candidate words and analogous words can be expected to slow the recognition process.

Minematsu neither discloses nor suggests applicant's claimed phrase matching. Nowhere does Minematsu reference a phrase, *per se*. The closest Minematsu comes to mentioning phrases is to vaguely refer to "*an index word including a plurality of words*" (column 11, line 26), "*a first candidate word including a plurality of words*" (column 12, line 42) and "*an analogous word including a plurality of words*" (column 12, line 48). Clearly, Minematsu considers phrases simply to be multiple words, does not recognize that they may have a pronunciation distinct from their component words (or syllables) and does not provide a distinct phrase database.

The above-referenced disclosures in Minematsu may at most be understood by one skilled in the art to suggest concatenation, or stringing together, of the pronunciations of the individual words or word parts, although no such explicit disclosure or suggestion is present in Minematsu. Even that suggestion may only be found by employing impermissible hindsight in the light of applicant's invention. The cited texts do not remotely suggest a database incorporating correctly pronounced phrases having pronunciations differing from the combined pronunciations of the individual word or word part components of the phrases.

The Office relies upon the disclosure at column 12 lines 18-28 of Minematsu to anticipate the phrase-modeling and phrase recognition aspects of applicant's claimed invention (Office Action, page 3, paragraph 2). However, in this disclosure, Minematsu merely addresses the case where a single word (e.g. "chicken") is erroneously

recognized as a plurality of words (e.g. "check in"). Thus, Minematsu does not meet the requirements of applicant's amended claim 40. Specifically, Minematsu does not disclose the comparing of the digital representation of an audible sound component of the input speech to the digital representations of properly spoken phrases in a phrase model database as defined in clause d) of amended claim 40.

Nor does Minematsu explicitly disclose a method which employs a properly pronounced phrase model database, *per se*, as well as a word model database. Minematsu merely refers generally to "an analogous word including a plurality of words". This vague reference might at most be understood to suggest that Minematsu's word database might include groups of words as well as individual words. It does not suggest that there should be two distinct databases, one containing words and the other containing phrases. Also nothing suggests that the words or syllables in the phrases should be properly pronounced phrases as opposed to a group of words that may be individually properly pronounced but which are improperly sounded as a phrase. Nor is there any hint in Minematsu, or elsewhere in the art that improved recognition and processing can be obtained by employing individual word and phrase databases and using properly pronounced phrases in the phrase database.

Prior to applicant's claimed invention, no one recognized that phrase matching as well as word matching could improve both the accuracy and the efficiency of speech recognition. Accuracy because phrases may be pronounced differently from the sum of the individual word pronunciations. Efficiency arising from recognizing words in groups and avoiding fruitless attempts to match unique phrase sounds with dictionary words.

The art contains no motivation, nor any reason, to employ phrase matching in a speech recognition method, as defined in applicant's claims. Prior to the present

invention, one of ordinary skill in the art would have understood phrases to be adequately recognized as though they were simply a plurality of words, as taught by Minematsu.

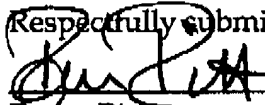
Accordingly base claim 40, as now amended is believed clearly and patentably distinguished from Minematsu or any other art known to applicant, and therefore allowable. System claim 57 has been comparably amended and is believed allowable for the reasons that claim 40 is believed allowable.

Claims 41-56 and 57-59 depend from base claims 40 and 57 respectively, and are therefore believed allowable with claims 40 and 57 for the reasons that claims 40 and 57 are believed allowable. Dependent claims 41-56 and 57-59 are furthermore believed clearly and patentably distinguished from the art of record, and therefore allowable, by the additional meaningful limitations they recite.

More particularly, Claim 42 specifically recites employing one or more speech-trained speakers to input phrases into phrase model databases, which is not remotely suggested by Minematsu or any of the other art of record in this application. By using high quality phrase models generated by skilled speakers, the subtleties of phrase pronunciation can more effectively be identified, enhancing recognition accuracy and speed.

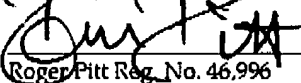
In contrast, applicant's invention, as claimed in amended claim 40 calls for sounds in the input speech to be compared, in the digital domain, with the properly spoken phrases and for a graphic representation, e.g. text, to be output when a match is made. This is not the use Minematsu discloses for his "analogous word including a plurality of words".

In view of the above amendments and the discussion relating thereto, it is respectfully submitted that the instant application, is in condition for allowance. Such action is most earnestly solicited. If for any reason the Examiner feels that consultation with Applicant's attorney would be helpful in the advancement of the prosecution, he is invited to call the telephone number below for an interview.

Respectfully submitted,  
By:   
Roger Pitt  
Reg. No. 46,996 Ph: (212) 536-4867

**Certificate of Transmission**

I hereby certify that this correspondence is being facsimile transmitted to the U.S. Patent and Trademark Office, to  
Fax No. (571) 273-8300, on December 28, 2005

  
Roger Pitt Reg. No. 46,996